2021 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS) | 978-1-6654-4067-7/21/$31.00 ©2021 IEEE | DOI: 10.1109/IEMTRONICS52119.2021.9422582

MLP for Spatio-Temporal Traffic Volume Forecasting

1st Asimina Dimara
Centre for Research and Technology Hellas
Information Technologies Institute
Thessaloniki, Greece, 57001
Email: adimara@iti.gr

2nd Dimitrios Triantafyllidis
Centre for Research and Technology Hellas
Information Technologies Institute
Thessaloniki, Greece, 57001
Email: triantafd@iti.gr

3rd Stelios Krinidis
Centre for Research and Technology Hellas
Information Technologies Institute
Thessaloniki, Greece, 57001
Email: krinidis@iti.gr

4th Konstantinos Kitsikoudis
Centre for Research and Technology Hellas
Information Technologies Institute
Thessaloniki, Greece, 57001
Email: kzkitsik@iti.gr

5th Dimosthenis Ioannidis
Centre for Research and Technology Hellas
Information Technologies Institute
Thessaloniki, Greece, 57001
Email: djoannid@iti.gr

6th Efthalia Valkouma
Egnatia Odos S.A
Thessaloniki, Greece
Email: thvalkou@egnatia.gr

7th Stilianos Skarvelakis
Egnatia Odos S.A
Thessaloniki, Greece
Email: sskarvel@egnatia.gr

8th Stavros Antipas
Patras Port Authority S.A
Patras, Greece, 26333
Email: santipas@patrasport.gr

9th Dimitrios Tzovaras
Centre for Research and Technology Hellas
Information Technologies Institute
Thessaloniki, Greece, 57001
Email: Dimitrios.Tzovaras@iti.gr

***Abstract*—Avoiding traffic congestion phenomena is an important aspect of efficient transportation infrastructure (e.g. toll roads) management. Traffic congestion phenomena can be avoided by forecasting volume traffic data. This paper aims to analyze how specific factors and parameters affect the behavior of traffic volume, like vehicle category, date, and weather data at a given timestamp and place. Moreover, the procedure of data prepossessing is presented to produce a cleaner data set that gives more fundamental information. Subsequently, spatio-temporal toll road prediction is achieved through a multi-layer perceptron. Finally, the proposed low-cost method is evaluated using real-life data from a toll plaza, while the experimental results show the efficiency of the proposed method.**

***Index Terms*—Traffic volume, spatio-temporal traffic, toll road traffic, traffic forecasting, data driven.**

I. Introduction

Traffic congestion not only leads to vehicular queuing and longer driving times but also results in waste fuel, increased carbon dioxide emissions and air pollution [1]. The first step towards solving traffic congestion is an integral approach which will enable traffic regulation through traffic volume management. Traffic management aims at safe travel and bottleneck-free roads and is part of the Intelligent Transportation Systems (ITS) [2]. A main aspect of ITS is to develop models that can be deployed to estimate, simulate and/or forecast traffic volume, traffic congestion and/or traffic density.

Especially, both the users and administrators of a toll road want to gain insight into the traffic volume at a specific toll plaza [3]. The existing literature referring to traffic forecasting is voluminous [4] and is implemented for many different aspects like travel time prediction, traffic volume prediction and waiting time in toll roads [5]. In [6] the authors state that using an artificial neural network (ANN) to predict traffic volume provides a lower error in comparison with other models (e.g. regression models). Furthermore, many of the existing models use cameras to detect traffic volume [7] and they are efficient but high-cost methods.

978-1-6654-4067-7/21/$31.00 ©2021 IEEE

Moreover, a fusion of ANN method and support vector machines is used to estimate traffic using many parameters like road vehicle counts, socio-economic data, geometry of the road and other parameters [8]. These type of models use a combination of road data and economic data to predict traffic [9] and are found to be accurate enough but harvesting all this type of data may prove a thorny task. Nonetheless, there is less literature about predicting volume traffic in toll plazas. Probably, because each local toll road is affected by multi-and various parameters like the local climate, distance from big cities, socio-economic data, and pavement quality. Therefore, even if a suggested method is accurate enough, it must be rebuilt for a specific toll plaza.

As a result, this paper suggests a low-cost method to forecast the traffic volume while using the least possible data that can be effortlessly retrieved like the local weather and number of vehicles passing daily through the toll road. Moreover, an effort is made to gain awareness of how this information affects volume traffic and employ the greatest possible advantage from it. Initially, the suggested methodology, data wrangling and preprocessing is addressed in Section II. Subsequently, the results are presented in Section III. Finally, conclusions are drawn in the final section.

II. Methodology

To achieve an increased performance some additional features that may affect the traffic volume have been integrated into the traffic forecasting using the proposed model [10]. In this paper, traffic data has been enhanced with temporal features (e.g., month, hours, day, and holiday periods) and weather features (e.g. temperature, cloud coverage and weather types (e.g sunny, foggy, rainy, snowy)). Specifically, temporal features are used to capture how seasonality may affect the traffic volume.

On the other hand, weather variables may not be directly linked to traffic volume forecasting, but they can affect it indirectly. Non expected weather conditions may disrupt holiday travel plans, while extreme weather conditions may lead to unexpected delays or postponement of business truck schedules. All the aforementioned circumstances that lead to schedule cancellations or postponements affect traffic, as a result necessitating the knowledge of weather conditions for more accurate predictions.

A. Data Wrangling

To explain the process with thoroughness, some further information about the data is provided. The vehicles are classified into four categories depending on the vehicle type as follows:

$$category = \begin{cases} CAT1, & \text{Motorcycle, Tricycle Vehicles} \\ CAT2, & \text{Passenger cars} \\ CAT3, & \text{Trucks, buses with} < 4 \text{ axes} \\ CAT4, & \text{Trucks, buses with} \geq 4 \text{ axes} \end{cases} \tag{1}$$

The toll road fees correspond to the four different vehicle categories, therefore a toll system can estimate the total number of vehicles for each category directly from the daily total sum of fees. As mentioned, the data set holds traffic information in a daily time series structure.

As it may be observed in Fig. 1 the number of vehicles that belong to CAT1 and CAT2 fluctuate similarly referring to months and are characterized by a traffic increase during the summer months (June, July, August). On the other hand, the number of vehicles that belong to CAT3 and CAT4 (Fig. 2) have similar patterns, with very small variations without following a certain motif during the year. This could indicate that CAT1 and CAT2 could be grouped together, as well as CAT3 and CAT4.

As a result, the problem could be split into 2 forecasting problems, forecasting of vehicles that belong to light vehicles (CAT1 and CAT2) and prediction of vehicles that belong to heavy vehicles (CAT3 and CAT4).

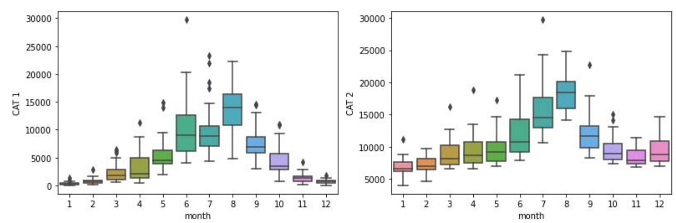

Fig. 1. CAT1, CAT2 per month

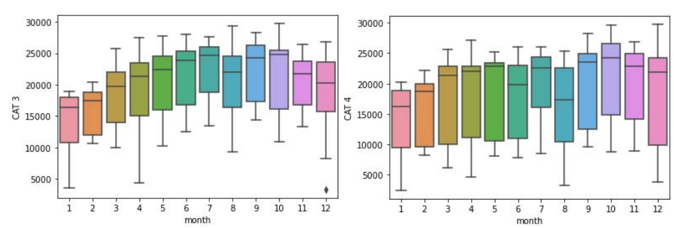

Fig. 2. CAT3, CAT4 per month

To further corroborate the previous assumption Fig. 3 and Fig. 4 reveals the same patterns in traffic volume for toll plazas for CAT1 and CAT2, and for CAT3 and CAT4. Specifically, on weekends more light vehicles seem to pass through the tolls than on weekdays, in contrast to heavy vehicles where more vehicles pass through the tolls on weekdays. Similar behaviors between the categories may also be noticed in Fig. 5 and Fig. 6.

It should be mentioned that for holiday periods the two grouped categories follow the same assumption as depicted in Fig. 7 and in Fig.8. Consequently, the problem may be split into light vehicles volume forecasting and heavy vehicles volume forecasting.

B. Data preprocessing

1) Impute missing values: The number of vehicles never have missing values as they derive from the fees. On the

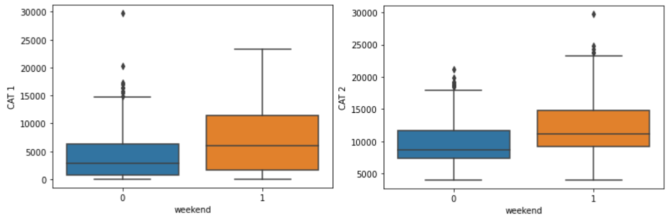

Fig. 3. CAT1, CAT2 per weekend

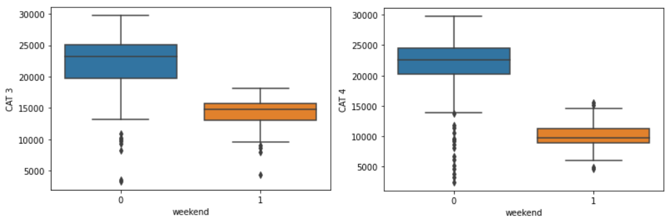

Fig. 4. CAT3, CAT4 per weekend

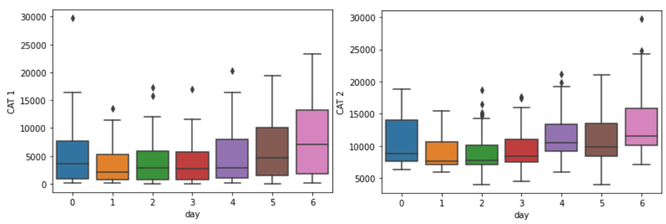

Fig. 5. CAT1, CAT2 per weekday

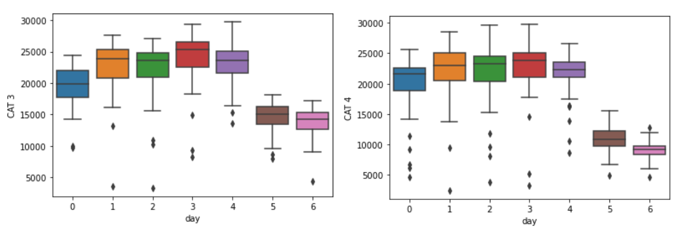

Fig. 6. CAT3, CAT4 per weekday

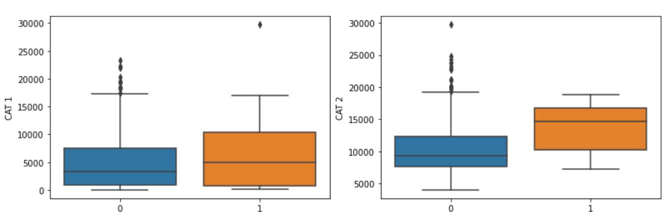

Fig. 7. CAT1, CAT2 per holidays

other hand missing data may appear in the weather variables. Every type of variable requires a different imputing method to be applied. For continuous values (e.g temperature), missing values are filled using the linear interpolation method [11], while missing data for categorical values (e.g weather state)

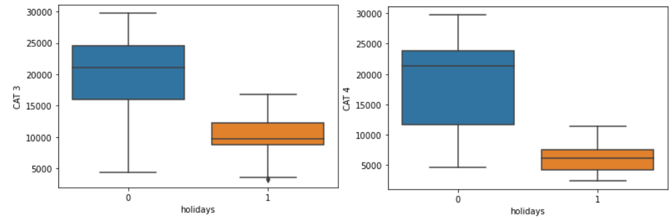

Fig. 8. CAT3, CAT4 per holidays

is replaced with the most frequent values within the data set [12].

2) Feature Encoding: Label encoding [13] is applied to the weather state feature to convert the possible weather states into a numeric form as described below:

$$weather\ state = \begin{cases} 0, & \text{Sunny} \\ 1, & \text{Foggy} \\ 2, & \text{Rainy} \\ 3, & \text{Snowy} \end{cases} \quad (2)$$

3) Outliers: Any outliers found were replaced by the average traffic of the current month for each category. In Figures 1 to 8 some outliers are observed to appear throughout the dataset. There are various methods to detect an outlier, however, within this paper, an empirical method for outlier dection was deployed, Z-score method [14], which describes the position of a record's distance (d) from the average (μ) measured in standard deviation unit (σ) and is defined as follows [14]:

$$Z_{score}(d) = \frac{d - \mu}{\sigma} \quad (3)$$

Any data that lie beyond 3σ from the average are considered as outliers and are replaced by the average traffic of the corresponding month.

4) Data modeling: To obtain insights about future traffic volume, past traffic data, temporal and weather variables were used. In addition, to achieve a better model performance, the method proposed in this paper also uses temporal and weather data from the nested period that are retrieved from Numerical Weather Prediction (NWP) models [15]. Specifically, for every record, data from n past days from current day (t) is selected and next day temporal and weather features in order to forecast the day ahead traffic volume ($t + 1$). The data modeling as described above is depicted in Fig. 9.

C. Traffic volume forecasting

Within this paper, a MLP model has been developed that implements the forecasting process, which is a conventional feed-forward Neural Network that contains more than one layer [16]. Some worth noting characteristics are that the first layer has as many nodes as the input features, while the number of the last layers' nodes depends on the machine learning task. For regression tasks, the last layer usually contains only one node. In the hidden or intermediate layers, the number of nodes may vary depending on the complexity

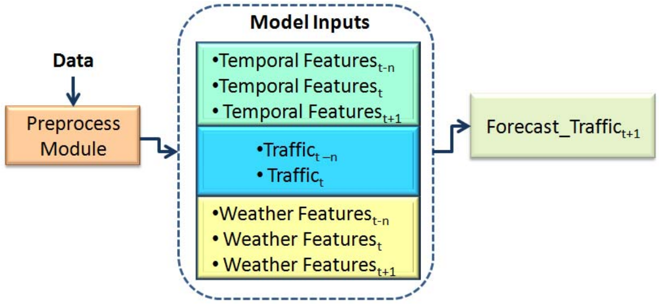

Fig. 9. Data transformation

of the problem or the source domain (e.g, images require a more complex architecture).

During training process, given a fully processed data set with N records, which can be defined as $\{(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_N, y_N)\}$, each $\mathbf{y}_i = \{y_1, \ldots, y_N\}$ is the ground truth number of vehicles and $\mathbf{x}_i = \{x_{i1}, \ldots, x_{im}\}$ represent the m input features as calculated in (4). As Fig. 9 implies traffic data from current and past n days are added as input. Temporal Features encapsulate 4 different features (month, day, weekend and holidays), while weather features contain 3 different features (temperature, cloud coverage, and weather state). For these contextual features, not only current and past n days are fed as input but also the day's ahead values that are retrieved from NPW models. To sum up, the number of input features a day's ahead prediction requires considering data from past n days can be calculated as follows:

$$\begin{aligned} Features = m &= (n+1) * 1[traffic\ data] \\ &+ (n+2) * 4[temporal\ features] \\ &+ (n+2) * 3[weather\ features] \\ &= 7 * (n+2) + n + 1 \\ &= 8 * n + 15 \end{aligned} \tag{4}$$

As the number of hidden layers and nodes per layer is increased the MLP becomes more complex and is more susceptible to overfitting. To measure the model's performance, a loss function is defined. The loss function the MLP Regressor uses to evaluate the model in each iteration is the mean squared error (MSE) [17]. The divergence between the predicted number of vehicles $\hat{y}_i$ and the actual number of vehicles y_i, calculated for N records is expressed with the following equation [17]:

$$Loss(y_i, \hat{y}_i) = \frac{1}{N} \cdot \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \tag{5}$$

Moreover, during the training procedure the weights are updated using an optimizer in order that mitigate the loss expressed by the loss function through an iterative process. In this work, Adam optimizer is used [18].

The MLP model is trained utilizing the aforementioned data. After the training phase the MLP forecasts the daily volume traffic. The retrieved data is pre-processed before fed in the MLP. Progressively, the MLP model is retrained every week with the new traffic and weather data in order to include the complete traffic volume and weather changes while improving itself through time or even capturing steep changes of the weather. A flow chart of the overall traffic forecasting model that was described is depicted in Fig. 10.

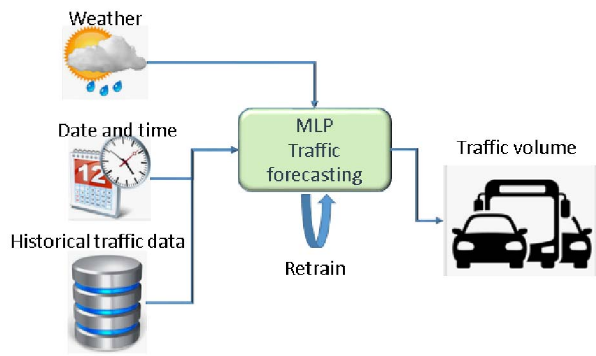

Fig. 10. Flow chart of the overall traffic forecasting model

III. Experimental Results

A. Use case

This section elaborates results obtained from the conducted experiments that took place at the toll plaza in Analipsi (Fig. 11), that is located between Kavala and Thessaloniki, Greece. The traffic data that were collected from the toll's implementation system were retrieved and used as input to the neural network that was developed within this work in order to forecast the day's ahead traffic volume. Furthermore, weather variables were retrieved based on the toll's exact location.

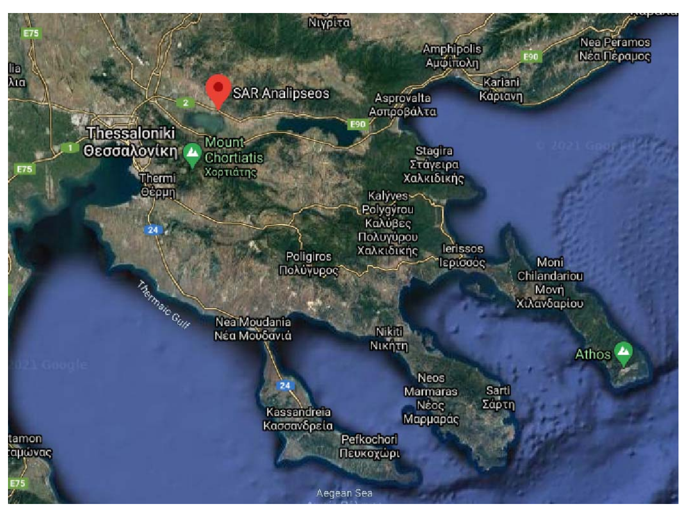

Fig. 11. Analipsi toll plaza map

The toll rates in Greece for various vehicles are as depicted in Fig. 12. It may be observed that the categories are the same as mentioned above for the CAT1, CAT2, CAT3, and CAT4, respectively.

Vehicle Categories		
1	Motorcycles	
2	Light vehicles with or without trailer with a height of up to 2.2m	
3	Vehicles with or without trailer, with two or three axes and a height more than 2.2m	
4	Vehicles with or without trailer, with four or more axes and a height more than 2.2m	

Fig. 12. Toll rate vehicle categories

B. MLP architecture

The lack of existence of a theory-based principle in order to define the optimal topology for a neural network often leads to trial and error methods for the definition of the network architecture [19]. However some broadly accurate practise-based guides exist that suggest the number of layers and neurons for a neural network [19]. As stated by Hecht Nielsen (1987) [20], the hidden layers' neurons should contain are:

$$nodes = 2m + 1, \tag{6}$$

where m is the number of input features (input nodes).

According to Huang [21] a neural network with two hidden layers should have a specific number of neurons. In the first hidden layer, the number of neurons that are suggested are:

$$nodes = \sqrt{(m+2)N} + 2\sqrt{\frac{N}{m+2}}, \tag{7}$$

while the sufficient number of neurons for the second hidden layer should be obtained from:

$$nodes = \sqrt{(m+2)N} \tag{8}$$

It is also worth mentioning that tuning the number of neurons using the power of two increases the speed of finding a layer size for a decent performance and leads to faster computational solutions [22]. Therefore, in the current work another empirical method was used defining the number of nodes as a power of two:

$$nodes = 2^k, \tag{9}$$

where $k \in \mathbb{N}$.

All above methods were tested for their performance through a trial and error procedure and the network with the highest performance was selected as the most appropriate. The training samples were approximately $N = 280$ while the input features considering the two previous days for the forecasting process are $m = 31$ (4).

$$m = 8 * n + 15 \overset{n=2}{=} 31 \tag{10}$$

According to Nielsen's method ([20] and (4)) each layer should contain $nodes = 63$, while using Huang's method [21] (Method 2) the first layer should contain $nodes \approx 100$, while the second layer should contain $nodes \approx 3000$.

The results obtained by the tuning procedure during the trial and error phase are presented in Fig. 13. Generally, the model seems to perform better having more than one layer, but utilizing Huang's method [21] (Method 2) the performance drastically decreases. As it may be observed in Fig. 13 Method 1 (Nielsen [20]) reveals poor results during all tests, while (9) shows better performance. Taken into account the overall performance of each network topology, method 3 (power of two [22]) with four hidden layers was selected as the most appropriate.

Method	Number of hidden layers	Number of neurons per layer	Validation R2 score
Method 1	1	63	0.68
	2	63-63	0.69
	3	63-63-63	0.76
	4	63-63-63-63	0.68
Method 2	1	100	0.76
	2	100-3000	0.78
	3	100-3000-3000	0.56
	4	100-3000-3000-3000	0.6
Method3	1	32	0.76
	2	64-128	0.85
	3	128-256-128	0.85
	4	**256-512-256-128**	**0.91**

Fig. 13. Tuning nodes per hidden layer

The overall MLP architecture that was selected is as depicted in Fig. 14.

C. MLP performance

In Fig. 15 and Fig. 16 the performance of both MLP models is depicted for heavy and light vehicles, respectively. The initial data set contains daily data during the period 2019 (January 1 - December 31) considering the total number of vehicles (for each category) that pass through the tolls. The data set was split to 75% for training and 25% for test. Specifically, the data samples that were used for training were selected by randomly getting three weeks from each month while the remaining week from each was used for testing. The metrics used to evaluate the MLP performance are Mean absolute Error (MAE) [23] and Mean Absolute Percentage Error (MAPE) [24]:

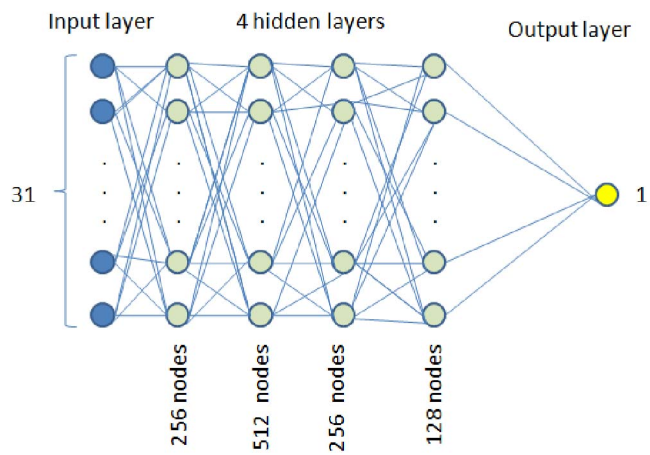

Fig. 14. MLP architecture

$$MAE = \frac{1}{N} \cdot \sum_{i=1}^{N} |y_i - \hat{y_i}| \quad (11)$$

$$MAPE = \frac{1}{N} \cdot \sum_{i=1}^{N} |\frac{y_i - \hat{y_i}}{y_i}| \quad (12)$$

An insubstantial deviation may be observed between the number of predicted and actual vehicles (Fig. 15 and Fig. 16). Furthermore, the predicted pattern seems to capture the actual fluctuations even in peak traffic volume moments. Specifically, the mean absolute percentage error in both cases is approximately 8.85%. Considering that 100 vehicles will pass through a specific toll plaza a MAPE score of this value implies that the forecasting model deviates by only 9 vehicles from the actual value.

In terms of performance, considering the nature of the data (total vehicles per day), the lack of intermediate values caused by the given time interval (aggregated daily values), is very reasonable and reveals decent credibility. To summarize, the selected architecture for the MLP proved to have an acceptable performance and is considered as appropriate for the forecasting process.

Conclusions

The proposed method is a tool for traffic volume forecasting at a specific toll plaza. Available data are analyzed revealing hidden information of the associated data, while the retrieved data is preprocessed to better fit the model. Traffic volume patterns during the experimental tests are accurate compared to the expected traffic volume.

The proposed model is a non-intrusive and low-cost traffic volume forecasting tool that informs users and administrators about the traffic volume of an exact location while indicating potential traffic congestion. Moreover, this traffic forecasting tool may be utilized in real-time traffic monitoring and management if data are retrieved real-time. Finally, it has the potential of enhancing the management of road networks while reducing traffic congestion.

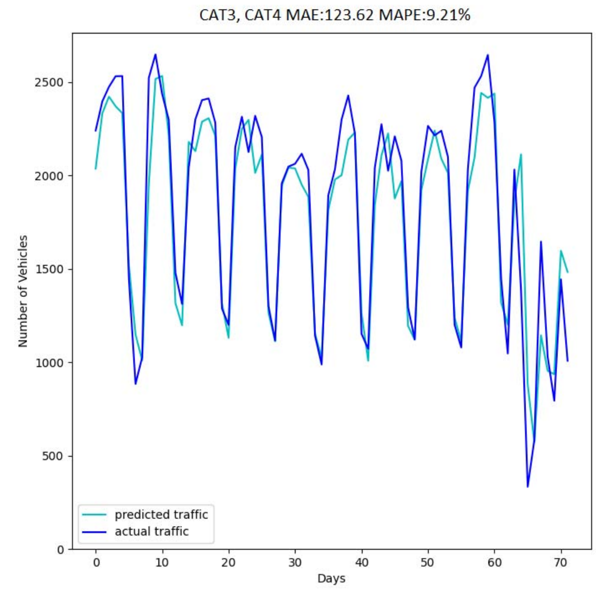

Fig. 15. Heavy Vehicles MLP performance

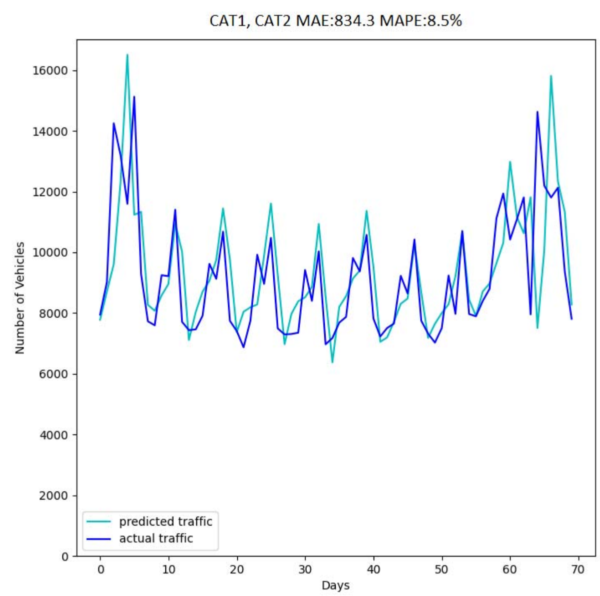

Fig. 16. Light Vehicles MLP performance

IV. Acknowledgments

This work is partially supported by the PERFFECT project co-financed by the European Regional Development Fund of the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH – CREATE – INNOVATE (project code:T1EDK-03368) and by the SELAS project co-financed by the European Regional Development Fund of the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH – CREATE – INNOVATE (project code:T1EDK-03547).

REFERENCES

[1] G. Li, W. Lai, X. Sui, X. Li, X. Qu, T. Zhang, and Y. Li, "Influence of traffic congestion on driver behavior in post-congestion driving," *Accident Analysis & Prevention*, vol. 141, p. 105508, 2020.

[2] W. Chen, L. Chen, Y. Xie, W. Cao, Y. Gao, and X. Feng, "Multi-range attentive bicomponent graph convolutional network for traffic forecasting," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, 2020, pp. 3529–3536.

[3] S. B. Seo, P. Yadav, and D. Singh, "Lora based architecture for smart town traffic management system," *Multimedia Tools and Applications*, pp. 1–16, 2020.

[4] G. Solvoll, T. A. Mathisen, and M. Welde, "Forecasting air traffic demand for major infrastructure changes," *Research in Transportation Economics*, vol. 82, p. 100873, 2020.

[5] D. Triantafyllou, N. Kotoulas, S. Krinidis, D. Ioannidis, and D. Tzovaras, "Large vehicle recognition and classification for traffic management and flow optimization in narrow roads," in *2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*. IEEE, 2017, pp. 1–4.

[6] R. K. Duraku and R. Ramadani, "Development of traffic volume forecasting using multiple regression analysis and artificial neural network," *Civil Engineering Journal*, vol. 5, no. 8, pp. 1698–1713, 2019.

[7] Z. Kadim, K. M. Johari, D. F. Samaon, Y. S. Li, and H. W. Hon, "Real-time deep-learning based traffic volume count for high-traffic urban arterial roads," in *2020 IEEE 10th Symposium on Computer Applications & Industrial Electronics (ISCAIE)*. IEEE, 2020, pp. 53–58.

[8] S. M. Khan, S. Islam, M. Z. Khan, K. Dey, M. Chowdhury, N. Huynh, and M. Torkjazi, "Development of statewide annual average daily traffic estimation model from short-term counts: A comparative study for south carolina," *Transportation Research Record*, vol. 2672, no. 43, pp. 55–64, 2018.

[9] M. Fu, J. A. Kelly, and J. P. Clinch, "Estimating annual average daily traffic and transport emissions for a national road network: A bottom-up methodology for both nationally-aggregated and spatially-disaggregated results," *Journal of Transport Geography*, vol. 58, pp. 186–195, 2017.

[16] J. Ilonen, J.-K. Kamarainen, and J. Lampinen, "Differential evolution training algorithm for feed-forward neural networks," *Neural Processing Letters*, vol. 17, no. 1, pp. 93–105, 2003.

[10] A. Dimara, D. Triantafyllidis, S. Krinidis, K. Kitsikoudis, D. Ioannidis, S. Antipas, and D. Tzovaras, "Fusing birch with g. boosting for improving temporal traffic congestion tailored to port gates: Case study in patras, greece," in *2020 IEEE 17th International Conference on Smart Communities: Improving Quality of Life Using ICT, IoT and AI (HONET)*. IEEE, 2020, pp. 1–5.

[11] Y. Yuan, C. Zhang, Y. Wang, C. Liu, J. Ji, and C. Feng, "Linear interpolation process and its influence on the secondary equipment in substations," in *2017 China International Electrical and Energy Conference (CIEEC)*, 2017, pp. 205–209.

[12] P. Schober and T. R. Vetter, "Missing data and imputation methods," *Anesthesia and analgesia*, vol. 131, no. 5, p. 1419, 2020.

[13] X. Liu, S. Wang, X. Zhang, X. You, J. Wu, and D. Dou, "Label-guided learning for text classification," *arXiv preprint arXiv:2002.10772*, 2020.

[14] P. J. Rousseeuw and M. Hubert, "Robust statistics for outlier detection," *Wiley interdisciplinary reviews: Data mining and knowledge discovery*, vol. 1, no. 1, pp. 73–79, 2011.

[15] D. Cho, C. Yoo, J. Im, and D.-H. Cha, "Comparative assessment of various machine learning-based bias correction methods for numerical weather prediction model forecasts of extreme air temperatures in urban areas," *Earth and Space Science*, vol. 7, no. 4, p. e2019EA000740, 2020.

[17] J. M. Martin-Donas, A. M. Gomez, J. A. Gonzalez, and A. M. Peinado, "A deep learning loss function based on the perceptual evaluation of the speech quality," *IEEE Signal processing letters*, vol. 25, no. 11, pp. 1680–1684, 2018.

[18] S. Bock and M. Weiß, "A proof of local convergence for the adam optimizer," in *2019 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2019, pp. 1–8.

[19] A. Dimara and C.-N. Anagnostopoulos, "Data based stock portfolio construction using computational intelligence," in *International Conference on Internet Science*. Springer, 2017, pp. 76–94.

[20] R. Hecht-Nielsen, "Kolmogorov's mapping neural network existence theorem," in *Proceedings of the international conference on Neural Networks*, vol. 3. IEEE Press New York, 1987, pp. 11–14.

[21] G.-B. Huang and H. A. Babri, "General approximation theorem on feedforward networks," in *Proceedings of ICICS, 1997 International Conference on Information, Communications and Signal Processing. Theme: Trends in Information Systems Engineering and Wireless Multimedia Communications (Cat.*, vol. 2. IEEE, 1997, pp. 698–702.

[22] R. K. Jain, "Rajeswari ponnuru (intel), ajit kumar p.(intel), and ravi keron n.(intel). cifar-10 classification using intel optimization for tensorflow," 2017.

[23] W. Wang and Y. Lu, "Analysis of the mean absolute error (mae) and the root mean square error (rmse) in assessing rounding model," in *IOP conference series: materials science and engineering*, vol. 324, no. 1. IOP Publishing, 2018, p. 012049.

[24] S. Kim and H. Kim, "A new metric of absolute percentage error for intermittent demand forecasts," *International Journal of Forecasting*, vol. 32, no. 3, pp. 669–679, 2016.